Four Essays on Applied Microeconometrics

Lopez Yurda, M.I.

2009

Chapter 6

Conclusions

This thesis is divided into two parts that could be read independently, covering topics in health economics and methods for longitudinal surveys.

In Chapter 2 we analyzed how macro-economic early-life conditions have a lasting effect on mortality and, specifically, how lower social classes suffer more from being born in recessions. We showed how being born into middle and higher class families, with higher average income, leads to a mitigation of the long-run health effects of adverse transitory macro shocks at birth. The results are not explained by social mobility, and social class and its interaction with transitory macro shocks at birth affect mortality at all ages. We found that social class effects are larger in absolute value for women than for men. This suggests that women's mortality is more strongly driven by the average economic situation in the family into which they are born, whereas the sensitivity for a negative shock, i.e. recession, at birth is similar, as is the extent to which the sensitivity for such a shock can be compensated for by social class. Having more comprehensive information about individual social mobility could be useful to provide insights into the differences in the results across gender. Besides, family tree data would have helped in providing a more thorough investigation of the issues we dealt with.

The results have some implications for policy design and policy analysis, notably in developing countries. Monitoring the health of babies and pregnant women among lower social classes in recessions might significantly reduce their mortality later in life. Projections of future health care and social security needs can be improved by taking long-run effects of early childhood business cycle conditions into account.

Chapter 2 provided a good setting for the next one. In Chapter 3 we focused on the business cycle early in life as a determinant of individual mortality later in life, performing a parallel study for eight European countries. To be able to perform an analysis of these characteristics we used cohort life-table data covering almost two centuries. We analyzed data from Denmark, England and Wales, Finland, Italy, The Netherlands, Norway, Sweden and Switzerland. The so-called Minimum Chi-Square method (Cockx, 1997) turned out to be an adequate tool for estimation. Though we found differences between countries, a majority of the results reflected a significant negative effect of the cyclical component of the business cycle at birth, implying that being born in a recession increases mortality later in life. As for the influence of the cycle during years of early childhood, results were less homogeneous. The effect was nonetheless negative and significant for many countries. The chapter demonstrated the possibility to perform a parallel analysis of considerable dimensions using a simple and computationally efficient methodology for data aggregated in life-tables.

The second part of this manuscript focused on the bivariate probit model with sample selection and the search for valid instruments to aid in its estimation. In Chapter 4 we examined, using Monte Carlo simulations, the performance of several parametric classical, parametric Bayesian and semi-nonparametric methods to detect and correct for sample selectivity in a censored probit model. We employed the Likelihood Ratio test to check the existence of selectivity. It is common in practice to use instruments to aid in identification of the bivariate probit model with sample selection. We thus performed a series of simulation experiments to study the case of continuous and dummy instruments, separately.

Results showed that the lack of informativeness of data from a bivariate probit model with sample selection can be exacerbated when the true correlation between the outcome and selection equation is high, and, even with as big a sample as $N = 5000$, detection of selectivity might not work satisfactorily when the Maximum Likelihood method is used for estimation. For a correctly specified model with continuous instruments, all procedures evaluated behaved in a somehow comparable way, in terms of bias of the coefficients in the outcome and selection equation. The time required for computations was disproportionate for the semi-nonparametric method, though. When dummy variables were introduced as instruments, estimation became more complicated. Of all the methods considered, Maximum Likelihood seemed to provide the best results, not only in terms of outcomes, but also in computational effort required. However, in the simulations the correlation coefficient is estimated with this approach at corner solutions in 38% of the cases when the true value of the parameter was high and as many as 100 dummies were used. Bayesian estimation fully avoided this problem, though the correlation coefficient estimate appeared to be strongly biased. The semi-nonparametric method gave a better outcome in this sense, though it still did not exceed Maximum Likelihood which, all in all, seemed to attain reasonably good results.

All the models estimated in this chapter were correctly specified. One question that arises is how already obtained results can change if the variables that are treated as instruments are excluded from the estimation. Besides, as solely the Likelihood Ratio test was used for selectivity testing, it would also be interesting to look at alternative tests and compare outcomes. Concerning the percentage of selection into the sample, we could as well replicate the same analyses in this chapter varying this number, to study, among other things, how this affects convergence rates and corner solutions in Maximum Likelihood estimation. As for the Bayesian procedure, Bayes factors can be used as a better and more reliable extension to the Likelihood Ratio test for the classical approaches. This could in future be incorporated to the empirical application.

Finally, Chapter 5 focused on the censored probit model previously studied to look for valid instruments to aid in its estimation. This was done in the context of longitudinal surveys that suffer from attrition that possibly does not completely occur at random. The choice of instruments usually depends on untestable assumptions. We used a linked survey and administrative database that allowed us to go beyond the traditional analyses in this field. The administrative records provided information on individuals labor market behavior and personal characteristics for the complete sample of participants and non-respondents. The survey was subject to attrition. We found that this attrition was selective for our outcome variable. We next explored candidate instruments based on interviewer information. As possible candidates we considered: the duration of the first interview, the number of interviews assigned to the interviewer in the first wave, the fraction of successfully completed interviews in the first wave and the identity of the interviewer who carried out the interviews (measured by interviewer dummies). The interviewer identity number and the interviewer success scores turned out to be valid instruments, meaning that they were informative on attrition behavior, but not related to the outcome variable of interest. These conditions did not hold for the other candidates. These results are of direct interest for agencies that run surveys as well as for researchers who are not so well endowed with data as in the chapter. We finally estimated the selection model with the interviewer success score as instrument in the attrition equation. The performance of this model was rather poor in the sense that the model had difficulties generating unconditional employment probabilities that were sufficiently close to probabilities based on a model that was estimated on the full sample (respondents and non-respondents). We argued that estimation of selection models with a binary outcome variable requires larger samples, as we explored in Chapter 4.